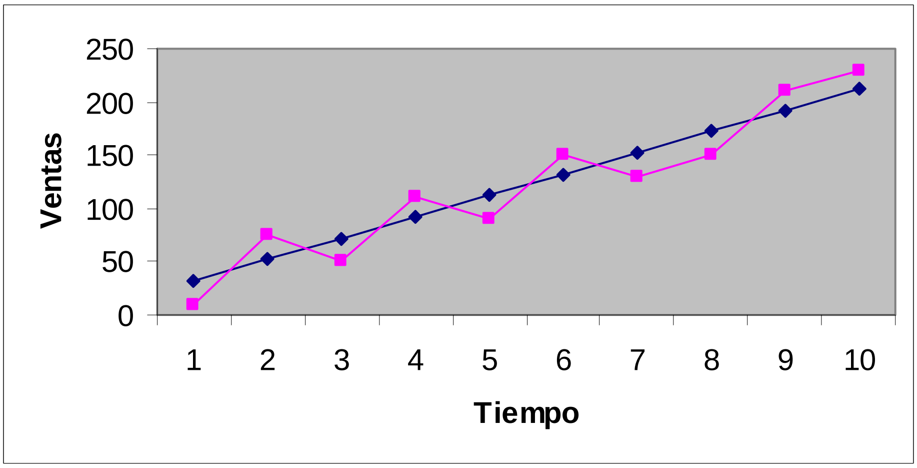
SERIES TEMPORALES
AJUSTE MÍNIMO CUADRÁTICO

1.- Modelo Lineal.

Supuesta una serie temporal de *n* términos (pares consecutivos de valores tiempo-ventas), en la que las ventas se hacen función del tiempo, de forma que $V = f(t)$, consideraremos la posibilidad de modelizar dicha relación mediante una expresión lineal en la que cada observación de la serie será el resultado de añadir, al valor que de dicha expresión lineal nos resulta para cada valor del tiempo, una variable residual que denominaremos error. Es decir que:

$$V_i = a + b\,t_i + e_i \qquad \forall \qquad 1 \le i \le n$$

Gráficamente podríamos representarlo por:



Mediante la aplicación del ajuste de regresión mínimo cuadrática, se buscan los valores de *a* y *b* que minimizan el sumatorio de los errores al cuadrado.

$$\text{Sea } a \text{ y } b \Rightarrow \sum_1^n e_i^2 \to Mínimo$$

Para ello, sabiendo que

$$e_i = V_i - a - bt_i \to \sum_1^n e_i^2 = \sum_1^n (V_i - a - bt_i)^2$$

Derivando respecto de *a* y de *b* e igualando a cero, se obtiene que los valores que cumplen con nuestro propósito de minimizar el sumatorio de los errores al cuadrado son los equivalentes a:

$$b = \frac{\sum_1^n V_i t_i - \bar{t}\sum_1^n V_i}{\sum_1^n t_i^2 - \bar{t}\sum_1^n t_i} \qquad \text{y} \qquad a = \bar{V} - b\bar{t}$$

Es capital para poder comprender el modelo que de éste método se obtiene, el hecho de entender que, si bien es posible obtener valores de a y b que lo representen, no por ello dicha representación debe ser siempre considerada como aceptable. Es fundamental establecer un índice que nos permita establecer si la recta obtenida por regresión representa adecuadamente a la serie temporal que modeliza. Para ello emplearemos el coeficiente de correlación lineal, cuya expresión es:

$$r = \frac{\sum_{1}^{n}\left(V_i - \bar{V}\right)\left(t_i - \bar{t}\right)}{\sqrt{\sum_{1}^{n}\left(V_i - \bar{V}\right)^2 \sum_{1}^{n}\left(t_i - \bar{t}\right)^2}}$$

Oscilando sus valores entre -1 y +1. Siendo más óptimo el modelo obtenido cuando más se acerca en valores absolutos el valor de *r* a 1, y menos cuando más cerca está de 0. Representando únicamente el signo que se obtenga la relación directa o inversa que existe entre las variable tiempo y venta.

2.- Extensiones al modelo lineal.

El examen detallado de una serie, gráficamente o analizando el valor del coeficiente de correlación lineal obtenido, puede llevarnos a la conclusión de que no es adecuado el ajuste lineal. Ante ésta situación es posible plantearse el ajustar la serie temporal a una función no lineal.

De entre todos los ajustes no lineales posible, es de destacar el <u>modelo potencial</u> y el <u>modelo exponencial</u>, cuya expresiones matemáticas son respectivamente:

$$V_i = \Omega\, t_i^{\Psi} \qquad \forall \qquad 1 \le i \le n$$

$$V_i = \Lambda\, \mathrm{E}^{t_i} \qquad \forall \qquad 1 \le i \le n$$

y que son fácilmente transformables en un modelo lineal mediante el empleo de logaritmos.

2.1.- Modelo Potencial.

De la expresión anterior del modelo potencial podemos obtener que:

$$\ln V_i = \ln \Omega + \Psi \ln t_i$$

De forma que si realizamos la siguiente transformación de variables

$$\Phi_i = \ln V_i$$
$$\Delta = \ln \Omega$$
$$\mathrm{T}_i = \ln t_i$$

Obtendremos una nueva expresión

$$\Phi_i = \Delta + \Psi \mathrm{T}_i \qquad \forall \qquad 1 \le i \le n$$

Que se corresponde con el modelo lineal en el apartado 1 expuesto, y del que consecuentemente se puede comprobar que:

$$\Psi = \frac{\sum_1^n \Phi_i \mathrm{T}_i - \bar{\mathrm{T}} \sum_1^n \Phi_i}{\sum_1^n \mathrm{T}_i^2 - \bar{\mathrm{T}} \sum_1^n \mathrm{T}_i} \qquad \text{y} \qquad \Delta = \bar{\Phi} - \Psi \bar{\mathrm{T}}$$

Siendo su correspondiente coeficiente de correlación lineal (entre las variables Φ_i y T_i que son las que se relacionan linealmente, y no para V_i y t_i cuya relación es potencial):

$$r = \frac{\sum_1^n \left(\Phi_i - \bar{\Phi}\right)\left(\mathrm{T}_i - \bar{\mathrm{T}}\right)}{\sqrt{\sum_1^n \left(\Phi_i - \bar{\Phi}\right)^2 \sum_1^n \left(\mathrm{T}_i - \bar{\mathrm{T}}\right)^2}}$$

No obstante, podemos conjuntamente medir la bondad del ajuste de nuestra serie al modelo potencial mediante el empleo del coeficiente de determinación general cuya expresión es:

$$R^2 = 1 - \frac{\sum_1^n e_i^2}{\sum_1^n \left(V_i - \bar{V}\right)^2}$$

Donde e_i se corresponde con la diferencia existente entre el valor de ventas observado en la serie y el que se obtiene para la misma referencia temporal con el modelo potencial obtenido.

$$e_i = V_i - \Omega\, t_i^{\Psi} \qquad \forall \qquad 1 \le i \le n$$

El coeficiente de determinación general tiene su rango entre 0 y 1, significando 0 que el modelo obtenido no representa en nada a la serie, y por el contrario 1 de que la representación es perfecta.

2.2.- Modelo Exponencial.

De la expresión anterior del modelo exponencial podemos obtener que:

$$\ln V_i = \ln \Lambda + t_i \ln \mathrm{E}$$

De forma que si realizamos la siguiente transformación de variables

$$\Phi_i = \ln V_i$$
$$\Theta = \ln \Lambda$$
$$\Gamma = \ln \mathrm{E}$$

Obtendremos una nueva expresión

$$\Phi_i = \Theta + \Gamma t_i \qquad \forall \qquad 1 \le i \le n$$

Que se corresponde con el modelo lineal en el apartado 1 expuesto, y del que consecuentemente se puede comprobar que:

$$\Gamma = \frac{\sum_1^n \Phi_i t_i - \bar{t} \sum_1^n \Phi_i}{\sum_1^n t_i^2 - \bar{t} \sum_1^n t_i} \text{ y } \qquad \Theta = \bar{\Phi} - \Gamma \bar{t}$$

Siendo su correspondiente coeficiente de correlación lineal (entre las variables Φ_i y t_i que son las que se relacionan linealmente, y no para V_i y t_i cuya relación es exponencial):

$$r = \frac{\sum_1^n \left(\Phi_i - \bar{\Phi} \right)\left(t_i - \bar{t} \right)}{\sqrt{\sum_1^n \left(\Phi_i - \bar{\Phi} \right)^2 \sum_1^n \left(t_i - \bar{t} \right)^2}}$$

No obstante, podemos igualmente medir la bondad del ajuste de nuestra serie al modelo potencial mediante el empleo del coeficiente de determinación general cuya expresión es:

$$R^2 = 1 - \frac{\sum_1^n e_i^2}{\sum_1^n \left(V_i - \bar{V} \right)^2}$$

Donde e_i se corresponde con la diferencia existente entre el valor de ventas observado en la serie y el que se obtiene para la misma referencia temporal con el modelo exponencial obtenido.

$$e_i = V_i - \Lambda \mathrm{E}^{t_i} \qquad \forall \qquad 1 \leq i \leq n$$

Siendo su rango entre 0 y 1, significando 0 que el modelo obtenido no representa en nada a la serie, y por el contrario 1 de que la representación es perfecta. Conviene hacer notar, que este cálculo del coeficiente de determinación general es válido siempre que el tipo de línea ajustada tenga una derivada parcial respecto del término no factor de la variable causal e independiente de la misma.

3.- La deflación de las series temporales.

La exposición ha sido desarrollada hasta éste punto sin considerar la diferencia y efectos que produce trabajar con ventas expresadas en términos de moneda nominal frente a moneda constante. Es evidente, que en las economías con niveles de precios no constantes (inflación o deflación) cualquier variable que se mida utilizando una misma moneda durante un periodo dilatado de tiempo produce una distorsión en la medición que se realiza al no estar empleando términos de evaluación homogéneos.

En consecuencia es fundamental el proceder a homogeneizar la serie temporal en cuestión refiriendo los términos de nuestras ventas a una cuantificación en términos de moneda constante mediante el empleo de los índices de precios al consumo.

Para el caso de España, los números índices de precios al consumo que permiten homogeneizar una serie de ventas, que se encuentra expresada en términos nominales, mediante su deflación, nos es ofrecida por el Instituto Nacional de Estadística (www.ine.es). Dichos índices se encuentran referenciados a un año base, y que en concreto para la estadística mencionada, según se recoge en sus indicaciones, son, los anteriores a 1.992 y los actuales a 2.001.

Conocido el mencionado índice, en tantos por uno, para un período de tiempo determinado, podremos homogeneizar las ventas en términos nominales de dicho período dividiéndolas por el expresado índice. Ventas que se encontrarán referidas en moneda constante del año base.

$$Ventas\,Cons\tan tes = \frac{Ventas\,No\min ales}{Indice\,\Pr ecio\,Consumo}$$

Para el caso de que se realice una previsión de ventas en función de modelos que se hayan basado en ventas medidas en términos constantes, y se quiera conocer dicha predicción en términos de moneda nominal, será necesario, multiplicar la previsión realizada en términos constante por el resultado de multiplicar el último índice conocido por la previsión de variación de precios -en tanto por uno- referida al periodo existente entre el último índice conocido y el periodo de referencia sobre el que se está realizando la previsión.

4.- Modelo de la función exponencial modificada.

La curva exponencial modificada tiene por ecuación

$$V_i = \alpha + \omega\, \theta^{\,t_i} \qquad \forall \qquad 1 \le i \le n$$

Debiéndose cumplir la condición de que $\alpha > 0$ y $0 < \theta < 1$.

Siendo *n* el número de periodos que contiene la serie, el método de cálculo del modelo exige que *n* sea múltiplo de 3. Si ello no es así se suprimen elementos seriales más antiguos (que son los de menor inercia), uno o dos como máximo, para que ésta condición se verifique. Haciendo $\lambda = n/3$.

A continuación, formándose las siguientes sumas:

$$S_1 = \sum_{1}^{\lambda} V_i = \lambda\alpha + \omega \sum_{1}^{\lambda} \theta^{t_i} = \lambda\alpha + \omega\theta \frac{\theta^{\lambda} - 1}{\theta - 1}$$
$$S_2 = \sum_{\lambda+1}^{2\lambda} V_i = \lambda\alpha + \omega \sum_{\lambda+1}^{2\lambda} \theta^{t_i} = \lambda\alpha + \omega^{\lambda+1}\theta \frac{\theta^{\lambda} - 1}{\theta - 1}$$
$$S_3 = \sum_{2\lambda+1}^{3\lambda} V_i = \lambda\alpha + \omega \sum_{2\lambda+1}^{3\lambda} \theta^{t_i} = \lambda\alpha + \omega^{2\lambda+1}\theta \frac{\theta^{\lambda} - 1}{\theta - 1}$$

se puede comprobar que:

$$\frac{S_3 - S_2}{S_2 - S_1} = \theta^{\,\lambda}$$

por lo que en consecuencia:

$$\theta = \sqrt[\lambda]{\frac{S_3 - S_2}{S_2 - S_1}}$$

Ahora, conocido θ, se puede hacer el siguiente cambio de variable

$$\chi_i = \theta^{\,t_i}$$

con lo que la ecuación primitiva queda:

$$V_i = \alpha + \omega\chi_i \qquad \forall \qquad 1 \le i \le n$$

en donde se puede estimar α y ω por el procedimiento expuesto en el apartado 1, ya que se trata de la ecuación de una recta. Pudiéndose comprobar que:

$$\omega = \frac{\sum_1^n V_i \chi_i - \bar{\chi} \sum_1^n V_i}{\sum_1^n \chi_i^2 - \bar{\chi} \sum_1^n \chi_i} \qquad \text{y} \qquad \alpha = \bar{V} - \omega \bar{\chi}$$

y donde, sus coeficientes de bondad serán los equivalentes a:

$$r = \frac{\sum_1^n \left(V_i - \bar{V} \right) \left(\chi_i - \bar{\chi} \right)}{\sqrt{\sum_1^n \left(V_i - \bar{V} \right)^2 \sum_1^n \left(\chi_i - \bar{\chi} \right)^2}}$$

$$R^2 = 1 - \frac{\sum_1^n e_i^2}{\sum_1^n \left(V_i - \bar{V} \right)^2}$$

Donde e_i se corresponde con la diferencia existente entre el valor de ventas observado en la serie y el que se obtiene para la misma referencia temporal con el modelo exponencial obtenido.

$$e_i = V_i - \alpha - \omega\, \theta^{t_i} \qquad \forall \qquad 1 \le i \le n$$

5.- El modelo logístico.

La función estándar del ciclo de vida de un producto se ajusta a una función en forma de S con un límite asintótico superior, que según el tipo de producto, mercado y otros condicionantes, presenta diversas variantes.

El modelo logístico clásico se basa sobre la hipótesis de que la tasa de aumento de las ventas por unidad de tiempo es proporcional a la parte del mercado que queda por saturar.

La expresión matemática de una de las posibles funciones logísticas que sirva de representación del ciclo de vida de un producto previo al desarrollo de su fase de declive es:

$$V_i = \frac{\delta}{1 + \rho e^{-\kappa t_i}} \qquad \forall \qquad 1 \le i \le n$$

donde e es el numero de Neper y δ, ρ y κ los parámetros a estimar para que quede definido el modelo logístico.

Trabajando con los valores recíprocos de la función anterior obtendremos:

$$\frac{1}{V_i} = \frac{1+\rho\, e^{-\kappa\, t_i}}{\delta} = \frac{1}{\delta} + \frac{\rho}{\delta} e^{-\kappa\, t_i} \qquad \forall \qquad 1 \le i \le n$$

y haciendo los siguientes cambios de variables:

$$\frac{1}{V_i} = \mathrm{H}_i \qquad \frac{1}{\delta} = A \qquad \frac{\rho}{\delta} = B \qquad e^{-\kappa} = C$$

la ecuación queda

$$\mathrm{H}_i = A + B\, C^{\,t_i} \qquad \forall \qquad 1 \le i \le n$$

que es la expresión de una exponencial modificada cuya estimación por el procedimiento de ajuste por regresión mínimo-cuadrático fue objeto de estudio en el apartado anterior.